

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/057,136
Source: JFW/6
Date Processed by STIC: 03/08/2006

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 03/08/2006

PATENT APPLICATION: US/10/057,136A

TIME: 11:34:34

Input Set : A:\Substitute Seq Text 47113.txt

Output Set: N:\CRF4\03082006\J057136A.raw

SEQUENCE LISTING

C--> 5 (1) GENERAL INFORMATION:

7 (i) APPLICANT: Therion, Corporation
 C--> 9 (ii) TITLE OF INVENTION: Recombinant Pox Virus For
 10 Immunization Against MUC1 Tumor-Associated Antigen
 12 (iii) NUMBER OF SEQUENCES: 21
 14 (iv) CORRESPONDENCE ADDRESS:
 15 (A) ADDRESSEE: NIXON PEABODY LLP
 16 (B) STREET: 100 Summer Street
 17 (C) CITY: Boston
 18 (D) STATE: MA
 19 (E) COUNTRY: USA
 20 (F) ZIP: 02110-2131
 22 (v) COMPUTER READABLE FORM:
 23 (A) MEDIUM TYPE: Diskette
 24 (B) COMPUTER: IBM Compatible
 25 (C) OPERATING SYSTEM: DOS
 26 (D) SOFTWARE: FastSEQ for Windows Version 2.0
 28 (vi) CURRENT APPLICATION DATA:
 C--> 29 (A) APPLICATION NUMBER: US/10/057,136A
 C--> 30 (B) FILING DATE: 25-Jan-2002
 W--> 41 (C) CLASSIFICATION:
 C--> 43 (vii) PRIOR APPLICATION DATA:
 34 (A) APPLICATION NUMBER: 09/366,670
 35 (B) FILING DATE: 03-AUG-1999
 39 (A) APPLICATION NUMBER: PCT/US98/03693
 40 (B) FILING DATE: 24-FEB-1998
 44 (A) APPLICATION NUMBER: 60/038,253
 45 (B) FILING DATE: 24-FEB-1997
 C--> 47 (viii) ATTORNEY/AGENT INFORMATION:
 48 (A) NAME: Eisenstein, Ronald I
 49 (B) REGISTRATION NUMBER: 30,628
 50 (C) REFERENCE/DOCKET NUMBER: 700953-047113-C2-RCE
 C--> 52 (ix) TELECOMMUNICATION INFORMATION:
 53 (A) TELEPHONE: 617-345-6054
 54 (B) TELEFAX: 617-345-1300
 55 (C) TELEX:
 58 (2) INFORMATION FOR SEQ ID NO: 1:
 60 (i) SEQUENCE CHARACTERISTICS:
 61 (A) LENGTH: 20 amino acids
 62 (B) TYPE: amino acid
 63 (C) STRANDEDNESS: single
 64 (D) TOPOLOGY: linear

RAW SEQUENCE LISTING

DATE: 03/08/2006

PATENT APPLICATION: US/10/057,136A

TIME: 11:34:34

Input Set : A:\Substitute Seq Text 47113.txt

Output Set: N:\CRF4\03082006\J057136A.raw

```

67      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
69  Gly Ser Thr Ala Pro Pro Ala His Gly Val Thr Ser Ala Pro Asp Thr
70    1          5          10          15
71  Arg Pro Ala Pro
72          20
74 (2) INFORMATION FOR SEQ ID NO: 2:
76      (i) SEQUENCE CHARACTERISTICS:
77          (A) LENGTH: 60 base pairs
78          (B) TYPE: nucleic acid
79          (C) STRANDEDNESS: single
80          (D) TOPOLOGY: linear
83      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
85  GGCTCCACCG CCCCCCAGC CCACGGTGTC ACCTCGGCC CGGACACCAG GCCGGCCCCG 60
87 (2) INFORMATION FOR SEQ ID NO: 3:
89      (i) SEQUENCE CHARACTERISTICS:
90          (A) LENGTH: 7 amino acids
91          (B) TYPE: amino acid
92          (C) STRANDEDNESS: single
93          (D) TOPOLOGY: linear
96      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
98  Pro Asp Thr Arg Pro Ala Pro
99    1          5
101 (2) INFORMATION FOR SEQ ID NO: 4:
103      (i) SEQUENCE CHARACTERISTICS:
104          (A) LENGTH: 60 base pairs
105          (B) TYPE: nucleic acid
106          (C) STRANDEDNESS: single
107          (D) TOPOLOGY: linear
110      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
112  GGCAGTACTG CACCACCGGC ACATGGCGTA ACATCAGCAC CTGATAACAAG ACCTGCACCT 60
114 (2) INFORMATION FOR SEQ ID NO: 5:
116      (i) SEQUENCE CHARACTERISTICS:
117          (A) LENGTH: 60 base pairs
118          (B) TYPE: nucleic acid
119          (C) STRANDEDNESS: single
120          (D) TOPOLOGY: linear
123      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
125  GGATCCACCG CGCCGCCTGC GCACGGAGTG ACGTCGGCGC CCGACACGCG CCCCCTCCC 60
127 (2) INFORMATION FOR SEQ ID NO: 6:
129      (i) SEQUENCE CHARACTERISTICS:
130          (A) LENGTH: 60 base pairs
131          (B) TYPE: nucleic acid
132          (C) STRANDEDNESS: single
133          (D) TOPOLOGY: linear
136      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
138  GGGTCAACAG CTCCTCCCGC TCATGGGGTT ACTTCTGCTC CAGATACTCG CCCAGCTCCA 60
140 (2) INFORMATION FOR SEQ ID NO: 7:
142      (i) SEQUENCE CHARACTERISTICS:
143          (A) LENGTH: 60 base pairs

```

RAW SEQUENCE LISTING

DATE: 03/08/2006

PATENT APPLICATION: US/10/057,136A

TIME: 11:34:34

Input Set : A:\Substitute Seq Text 47113.txt

Output Set: N:\CRF4\03082006\J057136A.raw

```

144         (B) TYPE: nucleic acid
145         (C) STRANDEDNESS: single
146         (D) TOPOLOGY: linear
149     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
151 GGTTCGACGG CCCCCCTGC TCACGGTGTA ACATCCGCCC CGGATACCAG ACCGGCCCCT      60
153 (2) INFORMATION FOR SEQ ID NO: 8:
155     (i) SEQUENCE CHARACTERISTICS:
156         (A) LENGTH: 60 base pairs
157         (B) TYPE: nucleic acid
158         (C) STRANDEDNESS: single
159         (D) TOPOLOGY: linear
162     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
164 GGCAGCACCG CACCGCCCGC ACACGGGGTC ACAAGCGCGC CAGACACTCG ACCTGCGCCA      60
166 (2) INFORMATION FOR SEQ ID NO: 9:
168     (i) SEQUENCE CHARACTERISTICS:
169         (A) LENGTH: 60 base pairs
170         (B) TYPE: nucleic acid
171         (C) STRANDEDNESS: single
172         (D) TOPOLOGY: linear
175     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
177 GGAAGTACCG CTCCACCTGC ACACGGGGTC ACAAGCGCGC CAGACACTCG ACCTGCGCCA      60
179 (2) INFORMATION FOR SEQ ID NO: 10:
181     (i) SEQUENCE CHARACTERISTICS:
182         (A) LENGTH: 60 base pairs
183         (B) TYPE: nucleic acid
184         (C) STRANDEDNESS: single
185         (D) TOPOLOGY: linear
188     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
190 GGGTCGACTG CCCCTCCGGC GCATGGTGTG ACCTCAGCTC CTGACACAAG GCCAGCCCCA      60
192 (2) INFORMATION FOR SEQ ID NO: 11:
194     (i) SEQUENCE CHARACTERISTICS:
195         (A) LENGTH: 60 base pairs
196         (B) TYPE: nucleic acid
197         (C) STRANDEDNESS: single
198         (D) TOPOLOGY: linear
201     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
203 GGTTCACCG CACCTCCAGC ACACGGAGTC ACGTCTGCAC CCGACACCCG TCCAGCTCCG      60
205 (2) INFORMATION FOR SEQ ID NO: 12:
207     (i) SEQUENCE CHARACTERISTICS:
208         (A) LENGTH: 60 base pairs
209         (B) TYPE: nucleic acid
210         (C) STRANDEDNESS: single
211         (D) TOPOLOGY: linear
214     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
216 GGTTAGTACAG CGCCACCCGC ACATGGCGTC ACGAGCGCTC CGGATACGAG ACCGGCGCCT      60
218 (2) INFORMATION FOR SEQ ID NO: 13:
220     (i) SEQUENCE CHARACTERISTICS:
221         (A) LENGTH: 78 base pairs
222         (B) TYPE: nucleic acid

```

RAW SEQUENCE LISTING

DATE: 03/08/2006

PATENT APPLICATION: US/10/057,136A

TIME: 11:34:34

Input Set : A:\Substitute Seq Text 47113.txt

Output Set: N:\CRF4\03082006\J057136A.raw

```

223         (C) STRANDEDNESS: single
224         (D) TOPOLOGY: linear
227         (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
229 GGCTCCACCG CACCCCAGC CCACGGTGTC ACCTCGGCCC CGGACACCAG GCGGGCCCCG 60
230 GGCTCCACCG CGGCCCCG                                     78
232 (2) INFORMATION FOR SEQ ID NO: 14:
234     (i) SEQUENCE CHARACTERISTICS:
235         (A) LENGTH: 60 base pairs
236         (B) TYPE: nucleic acid
237         (C) STRANDEDNESS: single
238         (D) TOPOLOGY: linear
241     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
243 GGCTCCACCG CCCCCCAGC CCATGGTGTC ACCTCGGCCC CGGACAACAG GCCCGCCTTG 60
245 (2) INFORMATION FOR SEQ ID NO: 15:
247     (i) SEQUENCE CHARACTERISTICS:
248         (A) LENGTH: 39 base pairs
249         (B) TYPE: nucleic acid
250         (C) STRANDEDNESS: single
251         (D) TOPOLOGY: linear
254     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:
256 GGCTCCACCG CCCCTCCAGT CCACAATGTC ACCTCGGCC 39
258 (2) INFORMATION FOR SEQ ID NO: 16:
260     (i) SEQUENCE CHARACTERISTICS:
261         (A) LENGTH: 20 amino acids
262         (B) TYPE: amino acid
263         (C) STRANDEDNESS: single
264         (D) TOPOLOGY: linear
267     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:
269 Gly Ser Thr Ala Pro Pro Ala His Gly Val Thr Ser Ala Pro Asp Thr
270  1           5           10           15
271 Arg Arg Ala Pro
272           20
274 (2) INFORMATION FOR SEQ ID NO: 17:
276     (i) SEQUENCE CHARACTERISTICS:
277         (A) LENGTH: 20 amino acids
278         (B) TYPE: amino acid
279         (C) STRANDEDNESS: single
280         (D) TOPOLOGY: linear
283     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:
285 Gly Ser Thr Ala Pro Pro Ala His Gly Val Thr Ser Ala Pro Asp Asn
286  1           5           10           15
287 Arg Pro Ala Leu
288           20
290 (2) INFORMATION FOR SEQ ID NO: 18:
292     (i) SEQUENCE CHARACTERISTICS:
293         (A) LENGTH: 13 amino acids
294         (B) TYPE: amino acid
295         (C) STRANDEDNESS: single
296         (D) TOPOLOGY: linear

```

RAW SEQUENCE LISTING

DATE: 03/08/2006

PATENT APPLICATION: US/10/057,136A

TIME: 11:34:34

Input Set : A:\Substitute Seq Text 47113.txt

Output Set: N:\CRF4\03082006\J057136A.raw

```

299      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:
301  Gly Ser Thr Ala Pro Pro Val His Asn Val Thr Ser Ala
302    1          5          10
304  (2) INFORMATION FOR SEQ ID NO: 19:
306      (i) SEQUENCE CHARACTERISTICS:
307          (A) LENGTH: 1527 base pairs
308          (B) TYPE: nucleic acid
309          (C) STRANDEDNESS: single
310          (D) TOPOLOGY: linear
312      (ix) FEATURE:
314          (A) NAME/KEY: Coding Sequence
315          (B) LOCATION: 1...1524
316          (D) OTHER INFORMATION:
318      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:
320  ATG ACA CCG GGC ACC CAG TCT CCT TTC TTC CTG CTG CTG CTC CTC ACA      48
321  Met Thr Pro Gly Thr Gln Ser Pro Phe Phe Leu Leu Leu Leu Leu Thr
322    1          5          10          15
324  GTG CTT ACA GCT ACC ACA GCC CCT AAA CCC GCA ACA GTT GTT ACG GGT      96
325  Val Leu Thr Ala Thr Thr Ala Pro Lys Pro Ala Thr Val Val Thr Gly
326          20          25          30
328  TCT GGT CAT GCA AGC TCT ACC CCA GGT GGA GAA AAG GAG ACT TCG GCT      144
329  Ser Gly His Ala Ser Ser Thr Pro Gly Gly Glu Lys Glu Thr Ser Ala
330          35          40          45
332  ACC CAG AGA AGT TCA GTG CCC AGC TCT ACT GAG AAG AAT GCT GTG AGT      192
333  Thr Gln Arg Ser Ser Val Pro Ser Ser Thr Glu Lys Asn Ala Val Ser
334    50          55          60
336  ATG ACA AGC TTG ATA TCG AAT TCC GGT GTC CGG GGC TCC ACC GCC CCC      240
337  Met Thr Ser Leu Ile Ser Asn Ser Gly Val Arg Gly Ser Thr Ala Pro
338  65          70          75          80
340  CCA GCC CAC GGT GTC ACC TCG GCC CCG GAC ACC AGG CCG GCC CCG GGC      288
341  Pro Ala His Gly Val Thr Ser Ala Pro Asp Thr Arg Pro Ala Pro Gly
342          85          90          95
344  TCC ACC GCC CCC CCA GCC CAC GGT GTC ACC TCG GCC CCG GAC ACC AGG      336
345  Ser Thr Ala Pro Pro Ala His Gly Val Thr Ser Ala Pro Asp Thr Arg
346          100          105          110
348  CCG GCC CCG GGC TCC ACC GCC CCC CCA GCC CAC GGT GTC ACC TCG GCC      384
349  Pro Ala Pro Gly Ser Thr Ala Pro Pro Ala His Gly Val Thr Ser Ala
350          115          120          125
352  CCG GAC ACC AGG CCG GCC CCG GGC TCC ACC GCA CCC CCA GCC CAC GGT      432
353  Pro Asp Thr Arg Pro Ala Pro Gly Ser Thr Ala Pro Pro Ala His Gly
354    130          135          140
356  GTC ACC TCG GCC CCG GAC ACC AGG CGG GCC CCG GGC TCC ACC CCG GCC      480
357  Val Thr Ser Ala Pro Asp Thr Arg Arg Ala Pro Gly Ser Thr Pro Ala
358  145          150          155          160
360  CCG GGC TCC ACC GCC CCC CCA GCC CAC GGT GTC ACC TCG GCC CCG GAC      528
361  Pro Gly Ser Thr Ala Pro Pro Ala His Gly Val Thr Ser Ala Pro Asp
362          165          170          175
364  ACC AGG CCG GCC CCG GGC TCC ACC GCC CCC CCA GCC CAT GGT GTC ACC      576
365  Thr Arg Pro Ala Pro Gly Ser Thr Ala Pro Pro Ala His Gly Val Thr

```

VERIFICATION SUMMARY

DATE: 03/08/2006

PATENT APPLICATION: US/10/057,136A

TIME: 11:34:35

Input Set : A:\Substitute Seq Text 47113.txt

Output Set: N:\CRF4\03082006\J057136A.raw

L:5 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]
L:9 M:220 C: Keyword misspelled or invalid format, [(ii) TITLE OF INVENTION:]
L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:36 M:238 W: Alpha Fields not Ordered, Reordered [(C) CLASSIFICATION:] of (1)(vi)
L:38 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:41 M:238 W: Alpha Fields not Ordered, Reordered [(C) CLASSIFICATION:] of (1)(vi)
L:43 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:47 M:220 C: Keyword misspelled or invalid format, [(viii) ATTORNEY/AGENT INFORMATION:]
L:52 M:220 C: Keyword misspelled or invalid format, [(ix) TELECOMMUNICATION INFORMATION:]

STATISTICS SUMMARY

DATE: 03/08/2006

PATENT APPLICATION: US/10/057,136A

TIME: 11:34:35

Input Set : A:\Substitute Seq Text 47113.txt

Output Set: N:\CRF4\03082006\J057136A.raw

Application Serial Number: US/10/057,136A

Alpha or Numeric or Xml: Alpha

Application Class:

Application File Date: 01-25-2002

Art Unit: IFW16

Software Application: FastSEQ2.0

Total Number of Sequences: 21

Total Nucleotides: 2304

Total Amino Acids: 614

Number of Errors: 0

Number of Warnings: 2

Number of Corrections: 2

MESSAGE SUMMARY

220 C: 2 (Keyword misspelled or invalid format)